

Bernhard Pörksen

## Das Menschenbild der Künstlichen Intelligenz

Ein Gespräch mit Joseph Weizenbaum

Joseph Weizenbaum, Jahrgang 1923, floh 1936 – als Kind jüdischer Eltern – vor den Nationalsozialisten in die USA und studierte später Mathematik. Als Assistent an der Wayne University in Detroit begann er, sich mit Computern zu beschäftigen, schrieb die Programmiersprache Slip, entwarf 1955 das Computersystem für die Bank of America und unterrichtete als Gastdozent an der Stanford-University. 1963 wurde Weizenbaum Professor für Computerwissenschaft am Massachusetts Institute of Technology (MIT), wo er bis zu seiner Emeritierung im Jahre 1988 lehrte. Berühmt hat in das Programm *Eliza* gemacht, das einen Psychotherapeuten simuliert, der sich – scheinbar zumindest – um psychologisches Verständnis seines Klienten müht und in seinen Antworten und Fragen an einen Psychotherapeuten der Schule von Carl Rogers erinnert. Seit dem Vietnamkrieg ist Weizenbaum in der Friedensbewegung aktiv; er gehört zu den Mitbegründern der „Computer Professionals for Social Responsibility“ und des „Forums der Informatiker für Frieden und gesellschaftliche Verantwortung“ in der Bundesrepublik. 1977 erschien sein Hauptwerk *Die Macht der Computer und die Ohnmacht der Vernunft* erstmals in deutscher Sprache: Es behandelt die Auswirkungen des Computers auf die menschliche Erfahrungswelt – damals ein neues und in seiner Brisanz noch weitgehend unbekanntes Thema. In zwei weiteren Büchern (*Kurs auf den Eisberg* und *Wer findet die Computermymen?*), die aus Gesprächen hervorgegangen sind, macht er auf die enge Verbindung von militärischer und computerwissenschaftlicher Forschung aufmerksam und kritisiert die Protagonisten der Künstlichen Intelligenz (KI) vehement.

### Das Modell der Maschine und das Modell des Menschen

**Pörksen:** Wer sich mit der Künstlichen Intelligenz befasst, der stößt immer wieder auf zwei grundsätzlich unterschiedliche Perspektiven: Manche KI-Protagonisten betrachten den Menschen als Modell der Maschine; die Maschinenintelligenz soll, so ihre Vorstellung, eine Imitation menschlicher Intelligenz darstellen. Bei einer anderen Gruppe von Forschern hat der Mensch diese Rolle als ideales Imitationsmodell jedoch eingebüßt: Jetzt sind es die Maschinen der Zukunft, die – wie man meint – dem Menschen überlegen sein werden. Sie erscheinen als weitgehend

vollkommene, mit einer phantastischen Denkkraft ausgestattete Entitäten. Läßt sich rekonstruieren, welcher dieser beiden Forschungsstränge am Anfang stand?

**Weizenbaum:** Diese beiden Stränge laufen seit den Anfängen der KI - der berühmten Dartmouth-Konferenz im Jahre 1956 - in irgendeiner Form nebeneinander her; auch schon zu Beginn kann man ahnen und sehen, dass nicht nur der Mensch als Vorbild und Ideal gilt, sondern dass man auch die Maschine als das Muster für das zu bauende Modell begreift. Auf der einen Seite findet sich also ein statement wie der Satz von Marvin Minsky, der gesagt hat, dass das Gehirn nichts weiter als eine *Fleischmaschine* sei. Im Original: „The brain is merely a meat machine.“ Man muss das wirklich auf englisch zitieren, denn im Englischen gibt es zwei Worte für Fleisch - *meat* und *flesh*. *Flesh* ist lebendes, ist lebendiges Fleisch. Mit *meat* kann man dagegen machen, was man will, man kann es essen, man kann es braten und verbrennen, es handelt sich um etwas Totes. In diesem Satz ist allerdings noch der Mensch das Muster; er ist es, dessen Intelligenz man nachbauen will. Was hier jedoch mitschwingt, ist die Idee, dass dies überhaupt keine Schwierigkeit sein dürfte; denn sein zentrales Organ - das Gehirn, das ihn determiniert und charakterisiert - sei eben bloß eine Fleischmaschine, nichts weiter. Auf der anderen Seite findet sich das ebenfalls von Marvin Minsky ausgesprochene Dogma, Gott sei kein besonders fähiger Ingenieur gewesen und bei dem Menschen handle es sich um eine Fehlentwicklung mit vielen überflüssigen Schwächen: Er braucht Schlaf, er lernt nur mühsam und vergißt wieder; außerdem ist er sterblich. Wir dagegen, so die in der KI-Gemeinde schon sehr früh verbreitete Auffassung, können etwas viel Besseres und Vollkommeneres erschaffen: Wir bauen das Idealmodell - eine Kreation, die sich zwar noch in manchem am Menschen orientiert, aber die der Perfektion doch näherkommt.

**Pörksen:** Das bedeutet - wenn man diesen Gedanken weiterspinnst - eigentlich, dass der Mensch aus dieser Perspektive zu einem evolutiven Bindeglied zwischen dem Affen und einem ihn übertrumpfenden Maschinenwesen wird. Luc Stells, KI-Professor an der Freien Universität Brüssel, schlägt vor, dieses Wesen der Zukunft „Homo cyber sapiens“ oder alternativ „Robot hominidus intelligens“ zu nennen.

**Weizenbaum:** In eine ähnliche Richtung zielt auch der Vorschlag, den Hans Moravec, Leiter des Mobile Robot Laboratory der Carnegie Mellon University, in seinem Buch *Mind children* macht: Er glaubt, das menschliche Bewusstsein ließe sich auf einen Roboter herunterladen; und man könne und solle es kopieren, um es von seinem notwendig sterblichen Körper abzulösen, der als bloßes Vehikel des Geistes gilt. Das Wesentliche des Menschen ist, so meint er, Information. Den Körper verachtet er.

**Pörksen:** Andere Autoren, die Moravec gedanklich nahestehen, nennen den menschlichen Körper ebenso wenig freundlich „ein Stück Fleisch“, „Brei“ oder „Sülze“.

**Weizenbaum:** So wird eben in dieser Szene geredet. In jedem Fall gelten die Roboter der Zukunft aus dieser Perspektive als eine Korrektur und Verbesserung der Natur: Moravec behauptet, dass solche Maschinen „unsere zivilisatorische Evolution vorantreiben“ könnten. Was er will, ist den bewussten und steuernden Eingriff in den Prozess der Evolution, den er für fehlerhaft hält. Seine Ideen gelten in den USA übrigens nicht als Science-fiction oder als bloße Trivialitäten, sondern sie kursieren an den Spitzenuniversitäten des Landes. Immerhin ist sein Buch von der Harvard University Press veröffentlicht worden.

### Prometheische Scham und prometheischer Stolz

**Pörksen:** In dieser Verehrung der Maschine als einer neuen Krone der Schöpfung drückt sich, so kann man vermuten, ein Grundgefühl aus, das der Philosoph Günther Anders einmal als „prometeische Scham“ bezeichnet hat: Er meint damit ein Gefühl der Unterlegenheit, so etwas wie Hilflosigkeit oder eben Scham angesichts der enormen Qualität der selbstentworfenen Maschinen. Teilen Sie die Diagnose von Günther Anders?

**Weizenbaum** Ganz gewiss gibt es dieses allgemeine psychologische Phänomen, dass man sich, wenn man mit einer besonders perfekten Leistung konfrontiert wird, unterlegen fühlt und sich für seine eigene Schwäche schämt. In diesem Sinne hat Günther Anders recht.

**Pörksen:** Marvin Minsky hat ja auch einmal gesagt, die Denkkraft der Silizium-Gehirne sei so phantastisch, dass wir – er meinte: die Menschen – froh sein könnten, wenn uns die Roboter der Zukunft als Haustiere behalten.

**Weizenbaum:** Aber es geht nicht allein um Scham: Beobachtbar ist eine Dialektik der Gefühle, die fast jeder Programmierer kennt. Ein gewisser Stolz bleibt auch dann bestehen, wenn man den Computer aus irgendeinem Grund für besser und klüger halten sollte. Denn immerhin war man selbst derjenige, der dieses besondere Programm geschrieben hat. Auf diese Weise entsteht ein grundsätzliches Paradox: Man schämt sich der eigenen Schwäche und ist doch gleichzeitig stolz, dass man ein derartig überlegenes Programm zu schreiben und zu erschaffen vermochte.

**Pörksen:** Moravec spricht auch von den neuen Robotern – das würde für diese These eines gewissen Stolzes sprechen – als seinen „Geisteskindern“.

**Weizenbaum:** Noch etwas anderes ist hier auffällig: Es ist ein Mann, der so spricht. Ist das Zufall? Ich glaube nicht, wenn man sich klarmacht, dass das Feld der KI von Männern dominiert wird. Mir scheint hier nicht nur der Wahn, Gott zu spielen, sondern auch der Neid auf die Frauen und ihre Fähigkeit, Kinder zu gebären, als ein treibendes Motiv. Was hier zum Ausdruck kommt, würde ich als Uterusneid bezeichnen.

**Pörksen:** Sie meinen ein komplementäres Phänomen zu dem von Sigmund Freud beschriebenen Penisneid?

**Weizenbaum:** Genau. Man tut nun so, als könne man auch Kinder hervorbringen – nur sind diese eben, wie man verbreitet, besser und intelligenter als jedes menschliche Wesen.

**Pörksen:** Aber derartige Behauptungen sind doch vollkommen spekulativ. Und auch die Ankündigung von Minsky, wir könnten froh sein, wenn uns die Maschinen der Zukunft „als Haustiere behalten“, scheint angesichts der Roboter, die man bis heute zustandegebracht hat, völlig überzogen. Könnte die relative Erfolgslosigkeit der KI-Forschung nicht gerade die Quelle eines neuen Selbstbewußtseins sein: Der Mensch sieht, was die Maschine alles nicht kann.

**Weizenbaum:** Natürlich könnte man, wenn man die unerfüllten Prognosen der KI betrachtet, sagen: Sind wir nicht wunderbar? Sind wir nicht großartig? Aber das ist Unsinn? Man sollte die menschliche Würde nicht danach bemessen, was eine Maschine kann oder nicht kann oder wie gut ein Computer Schach spielt. Der fatale Fehler dieser Form des noch humanistisch gemeinten Denkens ist, dass hier die Maschine zum Maß für den Menschen wird. Das Vermögen und Unvermögen eines Roboters ist doch keine Grundlage zur Bestimmung des menschlichen Selbstbewußtseins.

## Die heimliche Anthropologie der technischen Welt

**Pörksen:** Für mich stellt sich trotzdem die Frage, ob man die diversen Prognosen aus dem Fabelreich der KI eigentlich ernst nehmen sollte. Vieles scheint einfach absurd, manches ist aber auch witzig, amüsant und anregend: Man wird mit verrückten Wahrnehmungen, schrillen Ideen und verblüffenden Gedanken konfrontiert. Nochmals: Muß man die Vorstellungen eines Minsky oder Moravec ernst nehmen?

**Weizenbaum:** Das sind eigentlich zwei Fragen. Die erste lautet: Muß man damit rechnen, dass Menschen tatsächlich derartig überlegene Maschinen herstellen können? Müssen wir uns fürchten, dass derartige Roboter tatsächlich eines Tages existieren und wir in der Folge zu ihren Haustieren werden? Dazu sage ich: Nein, das müssen wir nicht. Und ich

füge hinzu, dass es auch wenig sinnvoll ist, mit diesen Leuten die Probleme der technischen Machbarkeit zu debattieren. Das ist ein *losing game*, ein verlorenes Spiel, das nur neue und vielleicht ebenso unerwünschte Entwicklungen anstößt. Die zweite und entscheidende Frage ist aber, ob man die Ideen, die hier verbreitet werden, ernstnehmen sollte. Und dazu sage ich: Ja, unbedingt! Der Grund ist, dass diese und andere Autoren an der Schwelle des Jahrtausends ein äußerst gefährliches Menschenbild propagieren.

**Pörksen:** Wie sieht dieses Menschenbild aus?

**Weizenbaum:** Es basiert auf der Vorstellung, der Mensch sei eine Maschine, die man im Prinzip und in naher Zukunft verstehen und entschlüsseln könne, um sie dann entsprechend zu korrigieren und zu verbessern. Das zentrale Dogma dieses Menschenbildes ist die Idee, dass jeder Aspekt des Lebens *computable* sei, dass er sich in berechenbare und formalisierbare Vorgänge auflösen ließe.

**Pörksen:** Wenn ich richtig verstehe, geht es Ihnen gar nicht primär um die Frage der technischen Realisierbarkeit, sondern um eine Art wortlose Ideologie, die mit der technischen Utopie ins Leben tritt. Marshall McLuhan hat wohl etwas sehr Ähnliches gemeint, als er von der *symbolischen Wirkung von Technik* – „the symbolic fallout of technology“ – sprach; Technik lasse, so McLuhan, ein Umfeld und eine Welt von Symbolen entstehen und verändere auf diesem Weg unser menschliches Selbstverständnis.

**Weizenbaum:** Sollte es tatsächlich möglich sein, derartig intelligente Maschinen herzustellen, dann wird dies erst in ferner Zukunft und nach vielen weiteren Generationen geschehen. Aber das Menschenbild wird, darauf kommt es an, heute geändert – und dies geschieht durchaus mit großem Erfolg: Minsky oder Moravec sind ja nicht einfach vom Himmel gefallen, sondern sie formulieren ihre Thesen in einer Kultur und in einer Zeit, in der man unbedingt an die Naturwissenschaften und die moderne Technologie glaubt.

**Pörksen:** Nochmals nachgefragt: Was sind die Gefahren eines Menschenbildes, das sich an Analogien aus der Welt der Maschinen orientiert?

**Weizenbaum:** Wir können aus der Geschichte dieses Jahrhunderts, vielleicht des brutalsten Jahrhunderts, lernen, welche entscheidende Rolle das Menschenbild in den Verbrechen der Vergangenheit spielte. Und wir müssen uns daran erinnern, dass die grausamsten Verbrechen möglich wurden, weil die Täter das Menschsein der Opfer leugneten. In der NS-Zeit stellte man Juden als Ungeziefer dar – eine Metapher, die den Massenmord legitimierte. Heute gewinnt, unterstützt durch die Autorität der Naturwissenschaften, die Vorstellung an Substanz und Macht, der

Mensch sei lediglich eine informationsverarbeitende Maschine, die von einem Roboter ersetzt werden könne. Es ist aus dieser Perspektive nur konsequent, dass Moravec ganz selbstverständlich von einer „postbiologischen Gesellschaft“ spricht, regiert von Robotern.

## Von den Folgen einer Metapher

**Pörksen:** Sie meinen, dass sich Tier- und Maschinenmetaphern ähneln?

**Weizenbaum:** Genau – und zwar in einer entscheidenden Hinsicht: Diese Metaphern vernichten die Ehrfurcht vor dem Menschen; sie lassen sein mögliches Ende erträglich erscheinen. Vor dem Ungeziefer und einer *meat machine* müssen wir keinen Respekt mehr haben. Man kann, wenn man von solchen Metaphern ausgeht, jede nur vorstellbare Grausamkeit verüben. Dies ist in unserem Jahrhundert auf furchtbare Weise bestätigt worden.

**Pörksen:** Sie haben insbesondere auf das Buch von Hans Moravec *Mind children* sehr wütend reagiert und gesagt, Moravec ziele letztlich auf eine „Endlösung der Menschheitsfrage“. Warum dieser Hinweis auf die Verbrechen der Nationalsozialisten? Mir leuchtet ein solcher Analogieverweis nicht wirklich ein.

**Weizenbaum:** Das mag sein. Aber wir sind alle das Ergebnis unserer Geschichte – einer Geschichte, die schon vor unserer Geburt beginnt. Was uns prägt, sind die Erfahrungen, die wir machen; was uns beeinflusst, ist die gesellschaftliche Situation, in die wir hineingeboren werden. Und so ist es nicht verwunderlich, dass gerade mir – vielleicht sollte ich in Anführungsstrichen hinzufügen: „einem deutschen Juden“ – diese Analogie einfällt. Ein anderer würde vermutlich etwas anderes bemerken.

**Pörksen:** Trotzdem stellt sich die Frage, worin Ihrer Auffassung nach die Gemeinsamkeit zwischen einem nationalsozialistischen Rassenwahn und den Phantasien eines KI-Professors besteht?

**Weizenbaum:** Natürlich stellt sich Hans Moravec nicht jeden Morgen vor den Spiegel und übt den Hitlergruß, das ist nicht der Punkt. Worauf ich hinauswill, ist die ungeheure Macht eines inhumanen Menschenbildes, das sich wie ein Virus in einer Gesellschaft auszubreiten vermag. Die wesentliche Gemeinsamkeit zwischen dem Nationalsozialismus und den Ideen eines Hans Moravec liegt, so meine ich, in der Entwürdigung des Humanen und der Phantasie eines perfekten neuen Menschen, der um jeden Preis geschaffen werden muß; am Schluß dieser Perfektionierung ist der Mensch allerdings nicht mehr da; er verschwindet in Moravecs

postbiologischer Gesellschaft. Ein humanes Menschenbild muß heute, davon bin ich zutiefst überzeugt, gegen die Dominanz der Maschinen-Metaphern verteidigt werden.

**Pörksen:** Tatsächlich ist ja niemand einem von der Maschinen-Metapher bestimmten Menschenbild hilflos ausgeliefert; Menschen empfinden doch sehr verschieden, sie agieren und rezipieren als Individuen, sind nicht notwendig gefügig gegenüber Wahrnehmungsrastern.

**Weizenbaum:** Das sehe ich anders; die Verletzbarkeit des Individuums ist groß. Sie lässt sich leicht durch die Bilder völliger Hingabe, die aus der NS-Zeit bekannt sind, belegen. Man denke nur an das Schreien und Jubeln und die ekstatischen „Sieg Heil“-Rufe der Deutschen. Menschen haben sich ja - gerade in diesem Jahrhundert - mehrfach begeistert ausgeliefert, bis es zu spät war und man entsetzt vor den Massengräbern stand.

### Das Credo der Künstlichen Intelligenz

**Pörksen:** Nun sind aber Metaphern aus der Welt der Maschine allgegenwärtig und waren immer schon da. Sie stammen noch aus den Zeiten der Dampfmaschinen und der Rädchen und Schrauben. Vom Herz reden wir als Pumpe; aus einer Sache, die wir betreiben, kann die Luft raus sein, eine andere läuft dagegen wie geschmiert. Von Menschen sagen wir, sie seien nicht mehr ganz dicht, sie würden durchdrehen, bei ihnen sei eine Schraube locker. Mit dem Jahrhundert der Elektrizität stehen wir dann plötzlich unter Strom und schalten nicht mehr schnell genug. Es brennen Sicherungen durch. Und so weiter.

**Weizenbaum:** Faktisch ist das mechanistische Welt- und Menschenbild, das uns solche Redewendungen beschert hat, etwa dreihundert Jahre alt. Insofern wäre ich sehr vorsichtig mit der Behauptung, derartige Metaphern seien immer schon da gewesen. Noch vor wenigen Jahrhunderten war für die Menschen Dantes Bild der Hölle und die Vorstellung eines Paradieses, in das man Dank eigener Anstrengungen oder der Güte Gottes gelangt, Realität. Aber natürlich stimme ich zu, wenn Sie sagen, dass Metaphern allgegenwärtig sind; jede Erkenntnis und jeder Versuch, etwas zu verstehen, sind - so würde ich sogar behaupten - letztlich metaphorischer Natur: Alles, was wir wissen, wissen wir in Form von Analogien, Vergleichen und Metaphern. Wenn wir glauben, dass wir etwas verstehen, so benutzen wir stets das Modell der Analogie, das heißt, wir konstruieren Ähnlichkeiten, um etwas Fremdes oder Unbekanntes im Rahmen von etwas Vertrautem und Bekanntem zu erfassen.

**Pörksen:** Könnte man dann nicht Marvin Minskys metaphorische Re-

de von der *meat machine* ähnlich undramatisch als einen Versuch begreifen, das Gehirn zu verstehen?

**Weizenbaum:** Natürlich, aber wenn ich sage, dass das Gehirn *nichts weiter als eine Fleischmaschine* ist, dann ändert sich etwas, und das ist entscheidend. Man meint nämlich, es handele sich um eine vollständige, umfassende und komplett ausreichende Beschreibung. Meine Kritik bezieht sich somit gar nicht auf die jeweilige Analogie oder die einzelne Metapher, die in irgendeinem Moment zur Beschreibung des Gehirns oder des Menschen verwendet wird. Dies ist in der Tat ein ganz gewöhnlicher und in der Wissenschaftsgeschichte spätestens seit Newton nachweisbarer Vorgang. Ich protestiere aber gegen die Behauptung, das Gehirn sei *nichts weiter als eine Fleischmaschine*, der Mensch sei nichts weiter als informationsverarbeitendes System.

**Pörksen:** Sie protestieren gegen Reduktionismus?

**Weizenbaum:** Mich entrüstet und ekelt das Credo der KI-Forschung, jeder Aspekt des menschlichen Lebens sei berechenbar und ließe sich entschlüsseln. Natürlich verarbeitet der Mensch Informationen. Und natürlich haben diese Metaphern als reduktionistische Modelle für die wissenschaftliche Arbeit und Erkenntnis ihren enormen Wert. Sie sind schlicht notwendig, sie sind brauchbar, um bestimmte Aspekte des menschlichen Lebens zu verstehen – aber sie sind eben, und dies gerät zunehmend aus dem Blick, Abstraktionen und erfassen nie das Ganze und Gesamte.

**Pörksen:** Welche andere Sprache oder Form der Beschreibung schwebt Ihnen vor?

**Weizenbaum:** Ganz spontan gesagt: Ich wäre – hätte ich tatsächlich die Autorität, derartige Vorschläge zu machen – in größter Versuchung, aus jedem wissenschaftlichen Text Formulierungen wie *nichts anderes als* ... zu streichen. Worauf es ankommt, ist, dass Phänomene und Objekte, Gefühle und Handlungen immer vielgestaltig sind und sich auf sehr verschiedene Weise charakterisieren lassen. Ihre Bedeutung wird erst in einem Kontext offenbar. Um ein einfaches Beispiel zu geben: Was bedeutet es, wenn man eine Hand auf der Schulter spürt? Eine Antwort ist nur möglich, wenn ich eine Geschichte erzähle. Nehmen wir einmal an, dass diese Geschichte von einem jungen Mann handelt. Er hat sich mit seiner Freundin gestritten und sitzt nun zerstreut und traurig in der Bibliothek und versucht zu arbeiten. Doch plötzlich spürt er eine Hand auf seiner Schulter. Oder nehmen wir an, dass es sich bei diesem jungen Mann um einen Verbrecher handelt: Er wird von der Polizei gesucht, sitzt im Wartesaal eines Bahnhofs – und plötzlich spürt er eine Hand auf der Schulter. Was ich damit sagen will? Nie ist etwas für immer und für alle Zeiten beschrieben; zu jeder Beschreibung gehört ein bestimmter Kontext.



## Vom Golem zum Roboter

**Pörksen:** Ich möchte – nach unserem kleinen Disput über notwendige und gefährliche Metaphern und Redewendungen – nochmals in anderer Weise versuchen, Marvin Minsky und Hans Moravec in Schutz zu nehmen. Man kann ja leicht erkennen, dass auch der Traum von einem künstlichen Menschen schon sehr alt ist. Moravec und Minsky erfüllen lediglich, so meine These, mit modernen Mitteln die Ziele des Rabbi Löw von Prag, der – glaubt man der Legende – einen Golem erschuf; sie folgen den Phantasien eines Paracelsus, der danach strebte den Homunculus zu schaffen; sie stehen in der Tradition der Automatenmanie, die sich bis in das Zeitalter der Aufklärung zurück verfolgen lässt.

**Weizenbaum:** Auch Ihre Frage enthält implizit die reduktionistische Formel *nichts anderes als ...* Sie sagen letztlich: Die Träume von Moravec und Minsky sind doch nichts anderes als der Traum, der schon die gesamte Geschichte der Menschheit hindurch geträumt wurde. Die Tatsache, dass die Pygmalion-Idee so alt ist und sich womöglich bis zu den Anfängen der Menschheit zurück verfolgen lässt, ist jedoch kein Beweis für die Harmlosigkeit von Moravec und Minsky, sondern belegt einfach nur die enorme Macht dieses Traums. Was sich zeigt, ist, dass Ideen, auch falsche Ideen, einen gewaltigen Einfluss besitzen. Der Unterschied zur Vergangenheit ist allerdings der gegenwärtige Glaube an die modernen Naturwissenschaften: Sie werden es uns, so meint man, in Kürze erlauben, den Menschen wirklich zu verstehen, und die Hochleistungscomputer erscheinen als die richtigen Werkzeuge, um den Pygmalion-Mythos, den man heute mit besonderer Euphorie und Arroganz vertritt, endlich Realität werden zu lassen.

**Pörksen:** Sind die Protagonisten der KI wirklich so arrogant und euphorisch? Zunehmend setzt sich doch die deutlich weniger reduktionistische Denkschule des Konnektionismus durch, die Intelligenz als Emergenzphänomen betrachtet: Man versucht, lernfähige neuronale Netze zu bauen, in denen Intelligenz plötzlich entsteht, emergiert. Das Gehirn betrachtet man aus dieser Perspektive als ein gigantisches neuronales Netzwerk, das auf der Basis zahlreicher Verknüpfungen arbeitet. Inspiriert durch die Frühzeit der Kybernetik und geschult an den Modellen der Selbstorganisation bekommt damit auch die mögliche Intelligenzleistung einer Maschine etwas Geheimnisvolles und unauflösbar Komplexes.

**Weizenbaum:** An eine neue Bescheidenheit der Konnektionisten glaube ich nicht. Auch Marvin Minsky ist inzwischen umgeschwenkt und hat erkannt, dass manche Ansätze aus den Anfängen der KI einfach falsch waren, selbst wenn man mit ihnen womöglich spektakuläre Ergebnisse erzielen konnte. Man musste sich eingestehen, dass der Kontext

eine zentrale Rolle spielt, wenn es um das Verstehen von Sprache geht. Man musste erkennen, dass der Körper des Menschen für unsere Art Intelligenz sehr entscheidend ist. Und es zeigte sich schließlich, dass es wenig Sinn hat, kleinere oder größere Programme herzustellen, die es der Maschine erlauben sollen, etwas scheinbar Intelligentes zu tun. Aber dieses Eingeständnis von Fehlern hat nicht zu einer anderen Haltung oder gar zu einer neuen Demut geführt. Im Gegenteil. Vielmehr wird die Einsicht in die eigenen Fehler als riesiger Fortschritt verkauft. Und die heimliche Hoffnung, den Menschen eines Tages zu verstehen und ein neuronales Netz, das dem Gehirn entspricht, zu bauen, um es dann in einem zweiten Schritt doch noch zu entschlüsseln, ist geblieben.

**Pörksen:** In Ihrer Kritik an der KI ist spürbar, dass Sie das Leben und den Menschen eigentlich für etwas fundamental Geheimnisvolles halten und sich wünschen, dass dieses Rätselhafte unserer Existenz anerkannt wird.

**Weizenbaum:** Die Welt ist voller Geheimnisse - und das Credo der KI-Szene, dass alles berechenbar sei, verleugnet das Geheimnis des Lebendigen, es erzeugt die Illusion vollständiger Durchschaubarkeit und legt nahe, alle Aspekte unserer Existenz seien enträtselbar. Der Glaube an Wunder und Mysterium erscheint aus dieser Perspektive lediglich als eine besondere Form von Dummheit. Mich verletzt diese Behauptung totaler Berechenbarkeit zutiefst.

## Mysterium und Geheimnis

**Pörksen:** Schon ziemlich am Schluß Ihres Buches *Die Macht der Computer und die Ohnmacht der Vernunft* findet sich eine Szene, die von dieser Begegnung mit dem Wunder und dem Geheimnisvollen handelt: „Als meine Kinder noch klein waren“, so liest man hier, „stand ich zuweilen zusammen mit meiner Frau über das Bett gebeugt, in dem sie schliefen. Wir sprachen miteinander, ohne zu reden; es war die Wiederholung einer Szene, so alt wie die Menschheit selbst. Es ist schon so, wie Ionesco einmal seinem Tagebuch anvertraut hat: „Man kann vieles mit Worten ausdrücken, nur nicht die lebendige Wahrheit.““

**Weizenbaum:** Meine Auffassung ist in der Tat, dass es etwas Unsagbares gibt, eine lebendige Wahrheit, die sich nicht in Worte fassen läßt.

**Pörksen:** Ein Mystiker würde in derselben Bedeutung von Gott sprechen.

**Weizenbaum:** Ich möchte eine kleine Geschichte erzählen. Am MIT arbeitete einmal ein episkopalischer Priester, er hieß kurioser Weise Scott Paradise. Eines Tages und schon Jahre nachdem wir uns kennengelernt und angefreundet hatten, feierte ich einmal eine kleine Party.

Scott Paradise war unter den Gästen. Meine Tochter Naomi sprach mit ihm und kam dann ganz erstaunt zu mir. Dein Vater, so hatte Scott Paradise zu ihr gesagt, ist ein besonders religiöser Mensch, er ist ein Mystiker.

**Pörksen:** Hat die Schärfe, mit der Sie gegen die KI-Forschung und die Computerisierung des Alltags streiten, demnach einen religiösen Halt?

**Weizenbaum:** Naja, es ist nicht meine Absicht, ein Glaubenssystem jüdisch-christlicher Herkunft oder eine organisierte Religion zu verteidigen; und ich denke auch nicht, dass es irgendwo oben im Himmel einen alten Mann gibt, der, in Bettwäsche gehüllt, das Geschehen in der Welt beobachtet, umgeben von blonden Frauen mit Flügeln, die um ihn herumfliegen. Aber es hat doch in meinem Leben die Erfahrung des Wunders gegeben. Da ist die Erfahrung der Trauer und der Erschütterung, da ist die plötzliche Freude am Morgen, da ist die Erfahrung der Liebe zwischen Menschen. Nun, ein Wunder ist eben ein Wunder; man kann es nicht beschreiben, man müsste ein Künstler sein, um sich ihm anzunähern.

**Pörksen:** Gott ist für Sie keine Person - was ist er dann?

**Weizenbaum:** Meine Antwort ist sehr einfach: Gott ist Liebe. Seine Allmacht und seine Gnade und die anderen Eigenschaften, die wir ihm zuschreiben, übersetze ich mir mit dem Wort Liebe. Ein solches Verständnis macht auch die Idee der göttlichen Allgegenwart auf einmal sinnvoll. Das bedeutet für mich nämlich, dass Gott auch in Auschwitz war, denn auch dort gab es Menschen, die versucht haben, aus Liebe zu helfen - und wenn diese Hilfe nur darin bestand, dass sie einem anderen die Hand auf die Schulter legten, um ihn zu trösten.

**Pörksen:** Gibt es Erfahrungen, autobiographische Erlebnisse, von denen Sie berichten mögen?

**Weizenbaum:** Als jüdische Kinder im Berlin der 30er Jahre das Gymnasium verlassen mussten, schickte man mich - ich war damals zwölf - auf eine jüdische Knabenschule. Dort hörte ich auf einmal Jiddisch und traf zum ersten Mal osteuropäische Juden, die im so genannten Scheunenviertel lebten. Sie waren entsetzlich arm. Ich habe mich damals regelrecht in einen jüdischen Jungen verliebt, der wirklich in Lumpen und Fetzen gekleidet war. Meine Mutter und ich haben versucht, ihm zu helfen, und ich habe ihm immer mal wieder ein paar Sachen mitgebracht. Es ist die Atmosphäre dieser Begegnung, die mich zutiefst gepägt hat. Sie lässt sich am Besten mit dem Wort *Liebe* beschreiben.

**Pörksen:** Die Vordenker eines postbiologischen Zeitalters würden sich vermutlich über einen derartigen Respekt vor dem Unsagbaren und dem Mysterium der Liebe lustig machen. So schreibt Max Moore, ein kalifornischer Autor und Vordenker der Roboter-Ära: „Religion ist eine Kraft, die unserer posthumanen Gesellschaft entgegensteht.“

**Weizenbaum:** Man macht sich lächerlich in bestimmten Kreisen, sicher. Man verbreite Opium für das Volk oder einfach Quatsch, so hört man. Zahlreiche moderne Menschen sind überzeugt, dass die Naturwissenschaft – im Gegensatz zu einer religiösen Idee – einen festen Boden des Wissens bietet. Ich bin da einer ganz anderen Auffassung. Für mich ist die Naturwissenschaft lediglich die heute vorherrschende Weltreligion mit Novizen (Studenten), Kirchen und Kathedralen (Universitäten), Priestern und Häretikern; auch die Kardinäle (die Nobelpreisträger) und ganz bestimmte Rituale, um das Ausmaß der eigenen Hingabe und Loyalität zu beweisen, fehlen nicht. Selbst die so genannten Laien glauben einfach – und zwar unabhängig von ihrer alltäglichen Erfahrung. Die meisten Menschen sind beispielsweise davon überzeugt, dass sich die Erde um die Sonne dreht und nicht umgekehrt. Sie können aber diese Annahme nicht beweisen und akzeptieren sie, obwohl sie ihrer alltäglichen Erfahrung widerspricht: Was sie sehen, ist eine bewegungslose Erde und eine sich bewegende Sonne. Und trotzdem vertrauen sie der Naturwissenschaft blind und unbegrenzt. Es ist diese quasi religiöse Inbrunst, die den Glauben unterstützt, die Naturwissenschaft werde es eines Tages möglich machen mit Hilfe der KI und der Genetik, einen künstlichen, einen besseren Menschen zu erschaffen. Und mit jedem neuen Triumph der Technik nimmt dieser Glaube weiter zu.

### Die unvermeidliche Individualität des Lebendigen

**Pörksen:** Es sind diese Triumphe der Technik und die sich heute bereits abzeichnenden Möglichkeiten der Menschenveränderung, die eine entschiedene Antwort auf die Grundfrage der philosophischen Anthropologie notwendig machen. Sie lautet: Was ist der Mensch?

**Weizenbaum:** Ich schlage vor, diese Frage etwas anders zu formulieren – sie könnte auch lauten: Was ist das Wesen des Menschen? Worin besteht die fundamentale Differenz zwischen dem Menschen und irgendeiner Maschine, die sich herstellen lässt? Und worin unterscheidet sich der Mensch von einem anderen beliebigen Wesen, das existiert? Es mag etwas arrogant klingen, aber meiner Meinung nach ist die Antwort äußerst offensichtlich und einfach: Ich werde nicht müde zu wiederholen, dass jeder Mensch das Ergebnis seiner persönlichen und individuellen Geschichte ist. Er hat seine eigene Biographie, er lebt zu einem bestimmten Zeitpunkt und ist unausweichlich in die ihn umgebende Gesellschaft eingebettet. Das heißt, jeder Mensch ist ein Sonderfall. Und wenn wir einmal hypothetisch die verrückte Idee für möglich halten, dass es eines Tages tatsächlich eine intelligente Maschine geben könnte, die sich wie

ein Mensch bewegt und ihm zum Verwechseln ähnlich sieht, dann wird diese doch niemals ein Mensch sein. Denn auch ihre Geschichte ist eine andere, sie muß eine andere sein - und sie wird sich deshalb immer von einem Menschen unterscheiden.

**Pörksen:** Zu rechnen ist wohl vorerst mit einem Mischwesen aus organischen und anorganischen Anteilen: Schon heute implantieren wir Elektronik in unsere Körper, benützen Herzschrittmacher und diverse Prothesen, tragen selbstverständlich Hörgeräte und lassen uns künstliche Hüften und Herzklappen einsetzen. Die Wissenschaftshistorikerin und Feministin Donna Haraway hat diese sich anbahnende Verschmelzung von Mensch und Maschine in ihrem berühmten *Manifest für Cyborgs* gefeiert und dazu aufgefordert, sie zu genießen und ihre emanzipative Kraft zu erkennen. Ihr feministisch inspiriertes Argument ist, etwas zugespitzt formuliert, dass durch Verschmelzung von Mensch und Maschine und die Entstehung des so genannten Cyborgs auch die Unterscheidung von Mann und Frau - das ist die fundamentale Differenz als Basis der Unterdrückung - ihre Grundlage verliert. Wenn wir alle endgültig zu Cyborgs werden, dann fehlt die Basis, um die Frau noch als die andere und als Unterdrückungsobjekt zu konstituieren. Haraway schreibt: „Daher steht die Cyborg-Politik auf dem Rauschen und auf der Verschmutzung und bejubelt die illegitime Verschmelzung von Tier und Maschine. Solche Verbindungen machen die Frau und den Mann problematisch, sie untergraben die Struktur des Begehrens, die imaginierte Macht, die Sprache und Gender hervorgebracht hat und unterlaufen damit die Strukturen und Reproduktionsweisen westlicher Identität, Natur und Kultur, Spiegel und Auge, Knecht und Herr, Körper und Geist.“

**Weizenbaum:** Das ist gut geschrieben, aber äußerst oberflächlich gedacht. Ein hohes Maß an Naivität, das stelle ich immer wieder fest, ist geradezu ein Merkmal der Tiefdenker, die sich mit Themen dieser Art beschäftigen. Was mich an den Ideen von Donna Haraway stört? Zum einen glaube ich überhaupt nicht, dass der Unterschied zwischen Mann und Frau aufgelöst wird. In den meisten Science-fiction-Filmen hat der Cyborg sehr wohl ein Geschlecht; und dieses Faktum sollte man nicht einfach, wenn man in dieser Weise argumentiert, ignorieren. Zum anderen ist diese kurze Passage, die Sie zitiert haben, ein Beispiel für eine andere Art der Fortschrittseuphorie: Man hofft auf die Maschinen als Hilfsmittel der Emanzipation, verlagert ein soziales Problem - den ewigen Konflikt zwischen Mann und Frau - ins Technische und behauptet, es besitze auch eine technische Lösung. Eine mathematische Gleichung kann ich gewiss lösen, aber menschliche und gesellschaftliche Probleme werde nie in dieser Weise endgültig und ein für allemal gelöst.

**Pörksen:** Aber müssen wir den Maschinen nicht in anderer Hinsicht

dankbar sein? Der Schriftsteller Peter Glaser wurde einmal von einem Fernsehsender dazu aufgefordert, in zehn Sätzen eine Rede an die Menschheit zu halten. Sein letzter Satz hat mich beeindruckt. Peter Glaser beendete seine kleine Rede mit der Aufforderung: „Seid freundlich zu den Maschinen!“ Das heißt für mich, dass die Maschinen womöglich nicht als Friedensstifter im Geschlechterkampf in Frage kommen, aber sie sind doch Erkenntnishilfen: Sie zwingen uns zur anthropologischen Reflexion und deshalb sollten wir ihnen dankbar sein und ihnen freundlich begegnen.

**Weizenbaum:** Auch mir gefällt dieser Satz, ich verstehe ich aber etwas anders. Wenn man seine Maschinen nicht verachtet, sondern sie gut behandelt, dann entsteht vielleicht auch eine andere Haltung gegenüber den Mitmenschen. Das ist nicht notwendiger Weise der Fall, aber immerhin möglich, und zumindest bedeutet dieser Satz, dass man überhaupt freundlich sein sollte – unabhängig davon, ob man es mit einer Maschine oder einem anderen Menschen zu tun hat. Dieser Schriftsteller sagt: „Seid freundlich zu den Maschinen!“ Man könnte auch sagen: „Seid freundlich zu der Welt!“

*Dieser Beitrag erscheint im Mai auch in dem von Bernd Flessner herausgegebenen Band: Nach dem Menschen. Der Mythos einer zweiten Schöpfung und das Entstehen einer posthumanen Kultur. Rombach-Verlag, Freiburg.*